

end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850. USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 192 a 262 c 251 g 170 t 2 others
ORIGIN

Query Match 34.1%; Score 874.6; DB 9; Length 877;
Best Local Similarity 99.7%; Pred. No. 5.2e-126;
Matches 874; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

29 CANGAGCCCAAGGCTGTCTGTCTGTGACAGAGCTCAAGAGGCTTGGGCTTCCCTCC 88
Db 1 CAAGAGCCCAAGGCTGTCTGTGACAGAGCTCAAGAGGCTTGGGCTTCCCTCC 60
Qy 89 TGGCTGGCTGTCTGTGAGAGGCTTCCAGTCCAGATCCCTAAGAGCATGGGAGC 148
Db 61 TGGCTGGCTGTCTGTGAGAGGCTTCCAGTCCAGATCCCTAAGAGCATGGGAGC 120
Qy 149 TGATCCATCCCTGTGTACAACTGTCTGTCTGTGACAGAGTGTGTAGCTACCCAAACAA 208
Db 121 TGATCCATCCCTGTGTACAACTGTCTGTCTGTGACAGAGTGTGTAGCTACCCAAACAA 180
Qy 209 CACCTAGCTCTCCCTGAGATCTCTCCAGGCTGAGAGATTTCTGGGTGTCTTAGACCA 268
Db 181 CACCTAGCTCTCCCTGAGATCTCTCCAGGCTGAGAGATTTCTGGGTGTCTTAGACCA 240
Qy 269 AGAGACTGGGAGATTTCCAGAAAGGCTTCAAGCTTCACTGTCCAGGCAAGATG 328
Db 241 AGAGACTGGGAGATTTCCAGAAAGGCTTCAAGCTTCACTGTCCAGGCAAGATG 300
Qy 329 CGTCTAGAGAGATGTCTTCCAGAGCTTGTGACAAACAAATTTCTCTGTATGT 388
Db 301 CGTCTAGAGAGATGTCTTCCAGAGCTTGTGACAAACAAATTTCTCTGTATGT 360
Qy 389 GCTTGAAGTGTCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448
Db 361 GCTTGAAGTGTCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 449 CAGGCCAAGCTTGAAGTGTCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
Db 421 CAGGCCAAGCTTGAAGTGTCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 509 GAGGCAAGGCTGAG 568
Db 481 GAGGCAAGGCTGAG 540
Qy 569 CGCTGAG 628
Db 541 CGCTGAG 600
Qy 629 TGTCTGAAGTGTCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688
Db 601 TGTCTGAAGTGTCTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 689 ATGGGTGGCTGTATAG 748
Db 661 ATGGGTGGCTGTATAG 720
Qy 749 GGAACCTTGAAG 808
Db 721 GGAACCTTGAAG 780
Qy 809 TGTCAAGTGGCTTGAAG 868
Db 781 TGTCAAGTGGCTTGAAG 840
Qy 869 GCCTGACAAATGGCTGTATAG 905

Db 841 GCCTGACAAATGGCTGTATAG 877

RESULT 2
B0054281 1020 bp mRNA linear EST 29-MAR-2002
LOCUS AGSCNCUR 6810234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
DEFINITION 5' mRNA sequence.

ACCESSION B0054281 GI:19813621
VERSION B0054281.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1020)
NIH-MGC http://mgi.nci.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLES Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/MCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LCM2125 row: j column: 11
High quality sequence stop: 555.
Location/Qualifiers

FEATURES
source 1. 1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/clone_11b="NIH_MGC_106"
/cissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pOT8; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 219 a 311 c 283 g 204 t 3 others
ORIGIN

Query Match 33.6%; Score 862.6; DB 14; Length 1020;
Best Local Similarity 95.5%; Pred. No. 3.4e-124;
Matches 973; Conservative 0; Mismatches 36; Indels 10; Gaps 8;

Qy 204 ACCAACAAGCTAGCTCTCCAG 263
Db 1 AACACACATAGCTCTCCAG 60
Qy 264 GACCAAG 323
Db 61 GACCAAG 120
Qy 324 GCATCGCTGAG 383
Db 121 GCATCGCTGAG 180
Qy 384 GATGCTCTTGAAGTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 443
Db 181 GATGCTCTTGAAGTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 444 TGTGCAAG 503
Db 241 TGTGCAAG 300

Qy	1033	ACTGTCGAGAGGACACACCTCACTAATGTGAAGAGCTGACAGCTCTCTGTGTTTCTGA	1098
Db	720	ACTGTCGAGAGGACACACCTCACTAATGTGAAGAGCTGACAGCTCTCTGTGTTTCTGA	779
Qy	1099	GCTGCGACA-GGGAGAGAGTCTCTTCTCAGTGAGAGG--TCCTCGGAGAGTCCCTCACTTC	1155
Db	780	ACTGCGACAGGGGAGAGAGTCTCTTCTCAGTGAGAGGATCCGGGAGAGTCCCTCACTTC	839
Qy	1156	TACATCAAGCTGTGAATGACGA-GGCTGTCTCTTGGAGATGTG-CTGAGGCCCAAGAGAG	1213
Db	840	TACATCAAGCCCTGTGATGACGAGGCTGTCTCTTGGAGATGCTCCAGCCCAAGAGAG	899
Qy	1214	GCCAAAAGGG--AAACCAAGCTGTGACACCT-ATAACCCCAATTACACCT--CTGAGCA	1268
Db	900	GCCCAAGGAGAAAACAGGCGCTGTGACCTTAAAAACCCCAATTAGCTTCTCTTGAGCC	959
Qy	1269	CCCCAGAGCGAAGCTGTGACACCTCAGAGAGGAGAGG	1304
Db	960	CCCCAAGCGAAGCTGTGAGCTTCTCCGAGGAGAAAAGG	995

BASE COUNT	195 a	200 c	236 g	219 t	4 others
ORIGIN					

Qy	1707	ATGATGCTTACCGAGAGACAACTTGGATGGTCCAGGCTCTATGTGACCTCCAGAGC	1766
Db	734	ATGATGCTTACCGAGAGACAACTTGGATGGTCCAGGCTCTATGTGACCTCCAGAGC	675
Qy	1767	AAAGAGAAAGACTTCGAGACAGCTAGAGTCTCAAAATGTCCCCCATTTAGAGCAAGAGCC	1826
Db	674	AAAGAGAAAGACTTCGAGACAGCTAGAGTCTCAAAATGTCCCCCATTTAGAGCAAGAGCC	615
Qy	1827	CAGCTCTTTTCTTTTCTTTTGGAGAGGAGCTTGCGCCCTGGTGGCCATGCTGGAGTGC	1886
Db	614	CAGCTCTTTTCTTTTCTTTTGGAGAGGAGCTTGCGCCCTGGTGGCCATGCTGGAGTGC	557
Qy	1887	TGGACGATCTCAGCTCACTGCAACCTCCTCATCTCTGGATTTCAAACTTCTCTGCTC	1946
Db	556	TGGACGATCTCAGCTCACTGCAACCTCCTCATCTCTGGATTTCAAACTTCTCTGCTC	497
Qy	1947	AGCCTCCAGAAATAGTGGAGATTACAGGGGTACACACATGTCAGGCTAATTTTGGTA	2006
Db	496	AGCCTCCAGAAATAGTGGAGATTACAGGGGTACACACATGTCAGGCTAATTTTGGTA	437
Qy	2007	TTTTTAGTACATGGGGTTTCCACCACTTGGCCAGGCTGGTGTGAACTTCTGACTCA	2066
Db	436	TTTTTAGTACATGGGGTTTCCACCACTTGGCCAGGCTGGTGTGAACTTCTGACTCA	377
Qy	2067	GGTATACCAACCACTTGGGCTCTCCAAAGTGTGGGATTACAGGATGAGACAGGACCC	2126
Db	376	GGTATACCAACCACTTGGGCTCTCCAAAGTGTGGGATTACAGGATGAGACAGGACCC	317
Qy	2127	CAGCTTAGCTCTCAGACTCTCAATTTCAATTTTGTGGCTTACCAATTCCTTACACTGACC	2186
Db	316	CAGCTTAGCTCTCAGACTCTCAATTTCAATTTTGTGGCTTACCAATTCCTTACACTGACC	257
Qy	2187	TTGGCATCTTGTGGCCGAAATAAATTAACCTCTTAAGTCTAGCAACTGACAGTGAAG	2246
Db	256	TTGGCATCTTGTGGCCGAAATAAATTAACCTCTTAAGTCTAGCAACTGACAGTGAAG	197
Qy	2247	CCAGGCACTCAGTGTCTGGCCAGGGCATCAGAAAGTGTGAAGCCCTCTCTCCACATGC	2306
Db	196	CCAGGCACTCAGTGTCTGGCCAGGGCATCAGAAAGTGTGAAGCCCTCTCTCCACATGC	137
Qy	2307	CAGAAGAGAGACACAGCTTACACCAATTCAGGCTTATTTCCCTGTGCTGCATTA	2366
Db	136	CAGAAGAGAGACACAGCTTACACCAATTCAGGCTTATTTCCCTGTGCTGCATTA	77
Qy	2367	ACAGAAAGAGTGTCTGTGATCCGCTAAGGATCAGGAGAGAGAAAGAGGAGTGGG	2426
Db	76	ACAGAAAGAGTGTCTGTGATCCGCTAAGGATCAGGAGAGAGAAAGAGGAGTGGG	17
Qy	2427	TGGAGAGACCCCTTC	2442
Db	16	TGGAGAGACCCCTTC	1

Db 660 CCCTTGATTCCTCTGCTG 677

RESULT 8 LOCUS

B0054265 986 bp mRNA linear EST 29-MAR-2002
AGENCY: 6830248 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936339
5', mRNA sequence.

ACCESSION B0054265.1 GI:19813605
VERSION
KEYWORDS
SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DSI/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM2125 row: 1 column: 12
High quality sequence stop: 515.
Location/Qualifiers
1..986
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5936339"
/clone_id="NIH_MGC_106"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pOT7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG (G) Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
SuperScript II RT (Life Technologies). Note: this is a
NIH MGC Library."

FEATURES

source

BASE COUNT 211 a 291 c 276 g 207 t 1 others
ORIGIN

Query Match 25.9%; Score 664.2; DB 14; Length 986;
Best Local Similarity 98.8%; Pred. No. 1.5e-93;
Matches 722; Conservative 0; Mismatches 3; Indels 6; Gaps 5;

QY 137 GCATGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACAGAGAGTGTGAGC 196
Db 1 GCATGGGAGCTGATCCATCCCTGGTGTACAACTGCTGACAGAGAGTGTGAGC 60
QY 197 TACCCAAACCAACAGCTGCTCCCTGGAAGTCTCCGAGGCTGAGAGATTTGGT 256
Db 61 TACCCAAACCAACAGCTGCTCCCTGGAAGTCTCCGAGGCTGAGAGATTTGGT 120
QY 257 GTCTTGAAGACCAAGACACTGGGAGACTTCCAGAAAGGCCCCCAAGCCCTTAACCTGTCC 316
Db 121 GTCTTGAAGACCAAGACACTGGGAGACTTCCAGAAAGGCCCCCAAGCCCTTAACCTGTCC 180
QY 317 AGCCAGAGCATGGCTCTCAGAGAGCTGTCTCCCAAGCCTTTGATGAGCAACCAATTTC 376
Db 181 AGCCAGAGCATGGCTCTCAGAGAGCTGTCTCCCAAGCCTTTGATGAGCAACCAATTTC 240
QY 377 CCTGATGATGCTCTTCTGAGTGTCTGCTGAGAAACAATGGAGAGTGTGCCCCAGAGAA 436
Db 241 CCTGATGATGCTCTTCTGAGTGTCTGCTGAGAAACAATGGAGAGTGTGCCCCAGAGAA 300

QY 437 GAAATCTGCGCAAGCCCAAGCTTGAAGTTCCTCTGTCCAAAGCCAGAGACCTGTACCA 496
Db 301 GAAATCTGCGCAAGCCCAAGCTTGAAGTTCCTCTGTCCAAAGCCAGAGACCTGTACCA 360
QY 497 TGAAGCAG 556
Db 361 TGAAGCAG 420
QY 557 CGGCGAGAGCTGTCTGAG 616
Db 421 CGGCGAGAGCTGTCTGAG 480
QY 617 GGTGAG 676
Db 481 GGTGAG 540
QY 677 CCAAGCTTCCAG 736
Db 541 CCAAGCTTCCAG 600
QY 737 TGTGTTACCTGAG 794
Db 601 TGTGTTACCTGAG 660
QY 795 AGGCTCTTACTCTCTCTGAGT-CCGCTTACAGCG-CCCTCATCTGAG--ACCGATCA 850
Db 661 AGGCTCTTACTCTCTCTGAGT-CCGCTTACAGCG-CCCTCATCTGAG--ACCGATCA 720
QY 851 GACACTACAGAG 861
Db 721 GACACTACAGAG 731

RESULT 9
LOCUS B6677567 794 bp mRNA linear EST 01-MAY-2001
DEFINITION B6677567.1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4748884 5',
mRNA sequence.
ACCESSION B6677567.1 GI:13908964
VERSION
KEYWORDS EST.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNL0601 row: d column: 05
High quality sequence stop: 790.
Location/Qualifiers
1..794
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4748884"
/clone_id="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NciI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT

192 a 229 c 212 g 161 t

ORIGIN

Query Match 25.5%; Score 653.8; DB 12; Length 794;
 Best Local Similarity 98.6%; Pred. No. 7, 2e-92;
 Matches 681; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

QY 930 GGTGACATTACTCTGAGCTGGCGATGACATCTGCTCCCTCACTCAAGAGCCCTGTGT 989
 DB 1 GGAGGTCTCTCTCAAGAGCTGGCGATGACATCTGCTCCCTCACTCAAGAGCCCTGTGT 60
 QY 990 CCTGAGAGGCTGGCGCTCCCTGCGAAGATATACCTCACTGAGCTGTGAGAG 1049
 DB 61 CCTGAGAGGCTGGCGCTCCCTGCGAAGATATACCTCACTGAGCTGTGAGAG 120
 QY 1050 GACACCACTCAACTGAGAAAGATGAGAGCTCCCTGCTTCTGAGAGCTGCAAG 1109
 DB 121 GACACCACTCAACTGAGAAAGATGAGAGCTCCCTGCTTCTGAGAGCTGCAAG 180
 QY 1110 GGAGAGTCTCTCTCAAGTGAAGGTCTCCGGAATCCCTCACTCAATCAAGCTGAA 1169
 DB 181 GGAGAGTCTCTCTCAAGTGAAGGTCTCCGGAATCCCTCACTCAATCAAGCTGAA 240
 QY 1170 TGACGAGGCTCTCTTGTGATGATGCTAGGCGCAAGAGAGGCGAAAGGAAACCA 1229
 DB 241 TGACGAGGCTCTCTTGTGATGATGCTAGGCGCAAGAGAGGCGAAAGGAAACCA 300
 QY 1230 AGGCTGCAACCTAGAAACCCCAATTCAGCTCTGCGCAACCCAGAGGCAAGCTGTGA 1289
 DB 301 AGGCTGCAACCTAGAAACCCCAATTCAGCTCTGCGCAACCCAGAGGCAAGCTGTGA 360
 QY 1290 CTCAGGAGAGGAGGTGGGACACAGAGGATCTAGAGGCCCACTGATCCCTGCTGT 1349
 DB 361 CTCAGGAGAGGAGGTGGGACACAGAGGATCTAGAGGCCCACTGATCCCTGCTGT 420
 QY 1350 TTCCTCTTGAAGCCCTTGAAGATCACTTCTTCAAGTCCATGATCCCACTGTGCA 1409
 DB 421 TTCCTCTTGAAGCCCTTGAAGATCACTTCTTCAAGTCCATGATCCCACTGTGCA 480
 QY 1410 CCTGATGAGAGTGCAGAGAGTGGGACAGAGGCGAGGCTTCCAAAAGATTAAGC 1469
 DB 481 CCTGATGAGAGTGCAGAGAGTGGGACAGAGGCGAGGCTTCCAAAAGATTAAGC 540
 QY 1470 CTCCTGGGGGGCTGACCTTGAAGTCTTGAAGTGGGGTTTCCAGTACCATGTGATG 1529
 DB 541 CTCCTGGGGGGCTGACCTTGAAGTCTTGAAGTGGGGTTTCCAGTACCATGTGATG 600
 QY 1530 CCTGCTGTTGAGAGGCTTCTTCAATCCCAAGATTAACAGGCGCCCAAGCAAGATA 1589
 DB 601 CCTGCTGTTGAGAGGCTTCTTCAATCCCAAGATTAACAGGCGCC--ACCAAGATA 658
 QY 1590 GAAACAAACCTTGAAGTCAACGAGAAATCA 1620
 DB 659 GAAACAA-CCCTGAGTCAACGAGAAATCA 688

RESULT 10 AL844308/c 642 bp mRNA linear EST 30-JUL-2002
 LOCUS AL844308

DEFINITION AL844308 pool_AK_11b_v_SPD Homo sapiens cDNA, mRNA sequence.

ACCESSION AL844308

VERSION AL844308.1 GI:22019090

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Bukariyola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 642)

Abneroff, K., Bethel, G., Bye, J. M., Howell, G. R., Huckle, E. J. and Sheridan, E.

REFERENCE Homo sapiens EST sequence

Unpublished (2002)

CONTACT: The Sanger Centre

FEATURES

source
 1..642
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="20"
 /clone_11b="pool_AK_11b_v_SPD"
 /note="Organ: breast; Vector: pZRo-1; Site 1: SphI;
 Site 2: SphI; Ductal carcinoma in situ, high-grade, comedo
 , from 41 yo female. Library constructed in the laboratory
 of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."
 pool_AK_11b_v_SPD cDNA library. Further information can be found at
 http://www.sanger.ac.uk/Teams/Team65/.

BASE COUNT

140 a 164 c 181 g 157 t

Query Match 24.6%; Score 630.8; DB 9; Length 642;
 Best Local Similarity 98.9%; Pred. No. 2, 9e-88;
 Matches 635; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1031 TACCTGTACTGTGACAGAGACACCACTCACTGAGAAAGCTGAGAGCTCCCTCTGT 1090
 DB 642 TACCTGTACTGTGACAGAGACACCACTCACTGAGAAAGCTGAGAGCTCCCTCTGT 583
 QY 1091 TTTCGAAAGCTGCACAGAGGAGAGGCTCTTCTCAAGTGAAGGTCTCCGGAAGCTCTCA 1150
 DB 582 TTTCGAAAGCTGCACAGAGGAGAGGCTCTTCTCAAGTGAAGGTCTCCGGAAGCTCTCA 523
 QY 1151 GCTTCAATCAAGCTGAATGACAGAGCTCTCTTGTGATGATGCTTACGCCCAAGAA 1210
 DB 522 GCTTCAATCAAGCTGAATGACAGAGCTCTCTTGTGATGATGCTTACGCCCAAGAA 463
 QY 1211 GAGGCGAAAGGAAACCAAGGCTGCACACTTGAAGACCCCAATTCAGCTCTGCGCAC 1270
 DB 462 GAGGCGAAAGGAAACCAAGGCTGCACACTTGAAGACCCCAATTCAGCTCTGCGCAC 403
 QY 1271 CCAGAGGCAAGCTGTGACCTGAGAGGAGGAGGTGGACACAGAGTGCATGAGGTC 1330
 DB 402 CCAGAGGCAAGCTGTGACCTGAGAGGAGGAGGTGGACACAGAGTGCATGAGGTC 343
 QY 1331 CCACCTGACCTTGTCTTCTCTCTCTTGTAGCCCTTGAAGTCACTTACTTCTTCAGT 1390
 DB 342 CCACCTGACCTTGTCTTCTCTCTCTTGTAGCCCTTGAAGTCACTTACTTCTTCAGT 283
 QY 1391 GCCATGATCCCACTTGCAGACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1450
 DB 282 GCCATGATCCCACTTGCAGACCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 223
 QY 1451 TTCCAAAAGAGATTAAGCTCTTGGGGGGTCTGACCTTGAAGTCTTGAAGTGGGGT 1510
 DB 222 TTCCAAAAGAGATTAAGCTCTTGGGGGGTCTGACCTTGAAGTCTTGAAGTGGGGT 163
 QY 1511 TTTCAGTACCATGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1570
 DB 162 TTTCAGTACCATGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 103
 QY 1571 AGGCCCCACCCAGAGTGAAGAAACACCCCTAGAGTCAACGAGAAAGCTTTTCAGAA 1630
 DB 102 AGGCCCCACCCAGAGTGAAGAAACACCCCTAGAGTCAACGAGAAAGCTTTTCAGAA 43
 QY 1631 ATCTACAGTCTGTTGAGACCAACCACTTCAAGAGGT 1672
 DB 42 ATCTACAGTCTGTTGAGAGCAACCACTTCAAGAGGT 1


```

RESULT 11
LOCUS      AL844312/c      619 bp      mRNA      linear      EST 30-JUL-2002
DEFINITION AL844312 pool_YT_11b_v_SPD Homo sapiens cDNA, mRNA sequence.
ACCESSION  AL844312
VERSION     AL844312.1 GI:22019094
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 619)
REFERENCE  Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and
AUTHORS   Sheridan,E.
TITLE     Homo sapiens EST sequence
JOURNAL   Unpublished (2002)
COMMENT   Contact: The Sanger Centre
            The Sanger Centre
            Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: humquery@sanger.ac.uk
            Sanger Centre name : sccdi0818.400489A
            Homo sapiens EST sequence. This sequence was generated as part of
            The Wellcome Trust Sanger Institute program to identify and
            annotate genes in the human genome. Incomplete or unconfirmed genes
            are experimentally analysed using a variety of cDNA library
            resources. This sequence was obtained from a PCR product generated
            from a pool of up to 100,000 cDNA clones derived from
            pool_YT_11b_v_SPD cDNA library. Further information can be found at
            http://www.sanger.ac.uk/Teams/Team69/.
            Location/Qualifiers
                1..619
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /map="20"
                /clone_lib="pool_YT_11b_v_SPD"
                /note="Organ: breast; Vector: pZE0-1; Site:1: Spn1;
                Site_2: Spn1; Ductal carcinoma in situ, high-grade, comedo
                , from 41 yo female. Library constructed in the laboratory
                of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."
```

BASE COUNT 133 a 159 c 177 g 150 t

ORIGIN

Query Match 24.0%; Score 615.8; DB:9; Length 619;
 Best Local Similarity 99.7%; Pred. No. 6.2e-86;
 Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1036 GTGACTGTGCGAGGACACCACTCACTGGAAGAAGCTGGAACAGCTCCCTCTGTTTCT 1095
DB 619 GTGACTGTGCGAGGACACCACTCACTGGAAGAAGCTGGAACAGCTCCCTCTGTTTCT 560
QY 1096 GAAGTCCACAGGGAGAGTCTCTCTAGTGAAGGTCTCCGGAGTCCCTCAGCTTC 1155
DB 559 GAGGTGCGACAGGAGGAGAGTCTCTCTAGTGAAGGTCTCCGGAGTCCCTCAGCTTC 500
QY 1156 TACATCAGCTGAATGACGAGGCTGTCTCTTTGATGATGCTGAGCCCAAGAGAGGC 1215
DB 499 TACATCAGCTTGAATGACGAGGCTGTCTCTTTGATGATGCTGAGCCCAAGAGAGGC 440
QY 1216 CAAAAGGAAACCAAGGCTGACACCTAGAACCCCAATTCAGCTCCTGCGGACCCCGA 1275
DB 439 CAAAAGGAAACCAAGGCTGACACCTAGAACCCCAATTCAGCTCCTGCGGACCCCGA 380
QY 1276 GGCAAGGCTGTGCACTCAGGAGGAGGAGTGAGACACAGAGTGCATCTAGAGTCCACC 1335
DB 379 GGCAAGGCTGTGCACTCAGGAGGAGGAGTGAGACACAGAGTGCATCTAGAGTCCACC 320
QY 1336 TGTACCTGTCTCTTCTCTCTTACCTTGAAGTCACTTCTCTTCAAGTGCAT 1395
DB 319 TGTACCTGTCTCTTCTCTCTTACCTTGAAGTCACTTCTCTTCAAGTGCAT 260
QY 1396 GATCCACCTGCGAAGCTTGTAGTGAAGTGAAGAAAGTGGAGACAGAGGCTTCA 1455
DB 259 GATCCACCTGCGAAGCTTGTAGTGAAGTGAAGAAAGTGGAGACAGAGGCTTCA 200
```

```

QY 1456 AANAAGATTAAGCTCTCTGAGGAGGCTGACCTAGTATGTTCTTGTGGGTTTCA 1515
DB 199 AANAAGATTAAGCTCTCTGAGGAGGCTGACCTAGTATGTTCTTGTGGGTTTCA 140
QY 1516 GTACCATCTGATGAGCCCTGCTGTGAGCCCATTCATCCCACTTAACAGAGCC 1575
DB 139 GTACCATCTGATGAGCCCTGCTGTGAGCCCATTCATCCCACTTAACAGAGCC 80
QY 1576 CCATCCCAAGATTAAGAAACCCCTTAGAGTCAACGAGAAAGTCTTTCAGAAATCTA 1635
DB 79 CCATCCCAAGATTAAGAAACCCCTTAGAGTCAACGAGAAAGTCTTTCAGAAATCTA 20
QY 1636 CAAGTCTGCTGAGAGCCAC 1654
DB 19 CAAGTCTGCTGAGAGCCAC 1
```

RESULT 12

LOCUS AL844307 614 bp mRNA linear EST 30-JUL-2002

DEFINITION AL844307 pool_AK_11b_v_SPD Homo sapiens cDNA, mRNA sequence.

ACCESSION AL844307

VERSION AL844307.1 GI:22019089

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 614)

REFERENCE Ashcroft,K., Bethel,G., Bye,J.M., Howell,G.R., Huckle,E.J. and

AUTHORS Sheridan,E.

TITLE Homo sapiens EST sequence

JOURNAL Unpublished (2002)

COMMENT Contact: The Sanger Centre

The Sanger Centre

Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

Email: humquery@sanger.ac.uk

Sanger Centre name : sccdi0818.154136A

Homo sapiens EST sequence. This sequence was generated as part of

The Wellcome Trust Sanger Institute program to identify and

annotate genes in the human genome. Incomplete or unconfirmed genes

are experimentally analysed using a variety of cDNA library

resources. This sequence was obtained from a PCR product generated

from a pool of up to 100,000 cDNA clones derived from

pool_AK_11b_v_SPD cDNA library. Further information can be found at

http://www.sanger.ac.uk/Teams/Team69/.

Location/Qualifiers

1..614

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="20"

/clone_lib="pool_AK_11b_v_SPD"

/note="Organ: breast; Vector: pZE0-1; Site:1: Spn1;

Site_2: Spn1; Ductal carcinoma in situ, high-grade, comedo

, from 41 yo female. Library constructed in the laboratory

of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."

BASE COUNT 134 a 188 c 164 g 128 t

ORIGIN

Query Match 23.7%; Score 609.2; DB:9; Length 614;
 Best Local Similarity 99.5%; Pred. No. 6.5e-85;
 Matches 611; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 882 CTGGCTGTACATCTACCGGCTCACTTCCCTCACTCAAGGCTGAGTCACTTA 941
DB 1 CTGGCTGTACATCTACCGGCTCACTTCCCTCACTCAAGGCTGAGTCACTTA 60
QY 942 CTGAGCTGGGAGTGAATCTGCTGCTTACCTCAAGAGCCCTGTCTTCAAGAGGC 1001
DB 61 CTGAGCTGGGAGTGAATCTGCTGCTTACCTCAAGAGCCCTGTCTTCAAGAGGC 120
QY 1002 TGGCCGCTCCCTGGAAGATTAACCTTACTGTATCTGACAGAGACCACTCA 1061
```

Db 121 TGAGCCGCTCCCTGGCAAGATATACCCCTACCTGTAGCTGACAGAGACACACACTCAA 180

Qy 1062 CTGGAAAGAGCTGACAGCTCCCTCTGTTTTCGAAAGCTGGCAGAGGGAGAGTCTCT 1121

Db 181 CTGGAAAGAGCTGACAGCTCCCTCTGTTTTCGAAAGCTGGCAGAGGGAGAGTCTCT 240

Qy 1122 TCTCAGTAGAGGCTCTCCGGAGATCCCTCACTTCTACATCAGCTGAATGACAGAGCTGT 1181

Db 241 TCTCAGTAGAGGCTCTCCGGAGATCCCTCACTTCTACATCAGCTGAATGACAGAGCTGT 300

Qy 1182 CTCTTTGATGATGCTTAGAGCCCAAGAGAGAGCCAAAGAGGAAACCAAGGCTGCACACC 1241

Db 301 CTCTTTGATGATGCTTAGAGCCCAAGAGAGAGCCAAAGAGGAAACCAAGGCTGCACACC 360

Qy 1242 TAGAACCCCAATTACGCTCTCTGGGACCCCAAGAGGAGAGCTGTGACATCAGAGAGGA 1301

Db 361 TAGAACCCCAATTACGCTCTCTGGGACCCCAAGAGGAGAGCTGTGACATCAGAGAGGA 420

Qy 1302 GGGTGGGACACAGAGGTGCATCTAGAGGTCCCACTGACCTGTGCTCTTCTCTTAG 1361

Db 421 GGGTGGGACACAGAGGTGCATCTAGAGGTCCCACTGACCTGTGCTCTTCTCTTAG 480

Qy 1362 CCCTTGAAGTACCTTACTCTCTTCCAGTGCATGATCCCACTGACCTGTGCTCTTAG 1421

Db 481 CCCTTGAAGTACCTTACTCTCTTCCAGTGCATGATCCCACTGACCTGTGCTCTTAG 540

Qy 1422 GTGCAGAGAGGTGGGACCAAGGSCCAGGGTTCGAAAGAGAAATAGGCTCTGGGGGT 1481

Db 541 GTGCAGAGAGGTGGGACCAAGGSCCAGGGTTCGAAAGAGAAATAGGCTCTGGGGGT 600

Qy 1482 CTGACCTAGTTAGT 1495

Db 601 CTGACCTAGTTAGT 614

RESULT 13

AL844310/c 611 bp mRNA linear EST 30-JUL-2002

LOCUS AL844310 pool_FLU_1lb_v_SPC Homo sapiens cDNA, mRNA sequence.

DEFINITION AL844310

ACCESSION AL844310

VERSION AL844310.1 GI:22019092

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 611)

Author(s) Aehroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and Sheridan, E.

REFERENCE Homo sapiens EST sequence

1 (bases 1 to 611)

Unpublished (2002)

CONTACT The Sanger Centre

THE SANGER CENTRE

Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

Email: humquerry@sanger.ac.uk

Sanger Centre name: sccid10817.4004894

Homo sapiens EST sequence. This sequence was generated as part of the Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool FLU_1lb v_SPC cDNA library. Further information can be found at <http://www.sanger.ac.uk/Teams/Teams9/>.

FEATURES

Source

1. 611

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="20"

/clone_1lb="pool_FLU_1lb v_SPC"

/notes="organ: breast; Vector: pBEO-1; Site: 1: SphI, Site_2: SphI; Ductal carcinoma in situ, high-grade, comedo

BASE COUNT 132 a 157 c 173 g 149 t

ORIGIN

Query Match 23.4%; Score 601.4; DB 9; Length 611;

Best Local Similarity 99.0%; Pred. No. 1.1e-83;

Matches 605; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1061 ACTGAAAGAGCTGACAGCTCCCTCTGTTTTCGAAAGCTGGCAGAGGGAGAGTCTC 1120

Db 611 ACTGAAAGAGCTGACAGCTCCCTCTGTTTTCGAAAGCTGGCAGAGGGAGAGTCTC 552

Qy 1121 TTCTCAGTAGAGGCTCTCCGGAGATCCCTCACTTCTACATCAGCTGAATGACAGAGCTGT 1180

Db 551 TTCTCAGTAGAGGCTCTCCGGAGATCCCTCACTTCTACATCAGCTGAATGACAGAGCTGT 492

Qy 1181 TCTCTTTGATGATGCTTAGAGCCCAAGAGAGAGCCAAAGAGGAAACCAAGGCTGCACACC 1240

Db 491 TCTCTTTGATGATGCTTAGAGCCCAAGAGAGAGCCAAAGAGGAAACCAAGGCTGCACACC 432

Qy 1241 CTAGAACCCCAATTACGCTCTCTGGGACCCCAAGAGGAGAGCTGTGACATCAGAGAGGG 1300

Db 431 CTAGAACCCCAATTACGCTCTCTGGGACCCCAAGAGGAGAGCTGTGACATCAGAGAGGG 372

Qy 1301 AGGTGGGACACAGAGGTGCATCTAGAGGTCCCACTGACCTGTGCTCTTCTCTTAG 1360

Db 371 AGGTGGGACACAGAGGTGCATCTAGAGGTCCCACTGACCTGTGCTCTTCTCTTAG 312

Qy 1361 GCCCTTGAAGTACCTTACTCTCTTCCAGTGCATGATCCCACTGACCTGTGCTCTTAG 1420

Db 311 GCCCTTGAAGTACCTTACTCTTCCAGTGCATGATCCCACTGACCTGTGCTCTTAG 252

Qy 1421 AGTGCAGAGAGGTGGGACCAAGGSCCAGGGTTCGAAAGAGAAATAGGCTCTGGGGGT 1480

Db 251 AGTGCAGAGAGGTGGGACCAAGGSCCAGGGTTCGAAAGAGAAATAGGCTCTGGGGGT 192

Qy 1481 TCTGACCTAGTTAGTCTTGAATTTGGGTTTTCAGTACATCTGATGCTGCTCTCTT 1540

Db 191 TCTGACCTAGTTAGTCTTGAATTTGGGTTTTCAGTACATCTGATGCTGCTCTCTT 132

Qy 1541 GAGCCCCATTCTACATCCCACTTAACAGAGGCCCAAGAGGTAGAAACCAAGCC 1600

Db 131 GAGCCCCATTCTACATCCCACTTAACAGAGGCCCAAGAGGTAGAAACCAAGCC 72

Qy 1601 TAGAGTCAACAGAAAGTCAATTTTCAGAAATCTACAGTCTGTGAGACACACAT 1660

Db 71 TAGAGTCAACAGAAAGTCAATTTTCAGAAATCTACAGTCTGTGAGAGACACAT 12

Qy 1661 ACCTCAGAAAG 1671

Db 11 ACCTCAGAAAG 1

RESULT 14

B0053486 878 bp mRNA linear EST 29-MAR-2002

LOCUS B0053486

DEFINITION AGENCOURT_6822017 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935253

ACCESSION B0053486

VERSION B0053486.1 GI:19812826

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 878)

Author(s) NIH-MGC <http://mgi.nci.nih.gov/>, National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: gsr@biml.nih.gov

Tissue Procurement: Dr. Daniel McVicar, DBS/NCI

CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LNCM2122 row: 1 column: 06
 High quality sequence stop: 394.

FEATURES
 source
 Location/Qualifiers

1. 878
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5935253"
 /clone_id="NIH MGC 106"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: blood; Vector: pOT7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

BASE COUNT 201 a 253 c 233 g 190 t 1 others

ORIGIN

Query Match 23.0%; Score 590.6; DB 14; Length 878;
 Best Local Similarity 95.5%; Pred. No. 3.9e-82;
 Matches 661; Conservative 0; Mismatches 25; Indels 6; Gaps 5;

142 GGGGAGCTGATCCATCCCTGGGTGACAACTGCTGACAGACAGATGCTGAGTACC 201
 Db 1 GGGGAGCTGATCCATCCCTGGGTGACAACTGCTGACAGACAGATGCTGAGTACC 59
 202 AAACCAACCTTACGCTCTCCCTGAGATCTCCAGGCTGAGAGATCTGGGTCTT 261
 Db 60 AAACCAACCTTACGCTCTCCCTGAGATCTCCAGGCTGAGAGATCTGGGTCTT 119
 262 AGACCAAGAGACCTGAGAGATCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAG 321
 Db 120 AGACCAAGAGACCTGAGAGATCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAG 179
 322 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 381
 Db 180 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 239
 382 ATGATGCTTCTGAGTCTCTGCTGAGAGACATGAGGAGGAGGAGGAGGAGGAG 441
 Db 240 ATGATGCTTCTGAGTCTCTGCTGAGAGACATGAGGAGGAGGAGGAGGAGGAG 299
 442 TCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 501
 Db 300 TCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 359
 502 GCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 561
 Db 360 GCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 419
 562 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
 Db 420 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479
 622 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
 Db 480 ACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539
 682 GTCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 741
 Db 540 GTCTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 597
 742 TTACCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 799

Db 598 TTACCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 657
 800 CTTACTCTCTGT-CAGTCCGCTCAGCCGCC 830
 Db 658 CTTACTCTCTGTGTAGTCCGCTTACGCCGCC 689

RESULT 15
 AL844309 606 bp mRNA linear EST 30-JUL-2002
 LOCUS AL844309 pool_Flu_11b_v_SPC Homo sapiens cDNA, mRNA sequence.
 DEFINITION AL844309
 ACCESSION AL844309.1 GI:22019091
 VERSION EST.
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 606)
 Ashcroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckie, E.J. and
 Sheridan, E.
 Homo sapiens EST sequence
 Unpublished (2002)
 CONTACT The Sanger Centre
 The Sanger Centre
 Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: humquery@sanger.ac.uk
 Sanger Centre name: sced10817.154136A
 Homo sapiens EST sequence. This sequence was generated as part of
 The Wellcome Trust Sanger Institute program to identify and
 annotate genes in the human genome. Incomplete or unconfirmed genes
 are experimentally analysed using a variety of cDNA library
 resources. This sequence was obtained from a PCR product generated
 from a pool of up to 100,000 cDNA clones derived from
 pool_Flu_11b_v_SPC cDNA library. Further information can be found
 at <http://www.sanger.ac.uk/Teams/Team69/>.

FEATURES
 source
 Location/Qualifiers

1. 606
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="20"
 /clone_id="pool_Flu_11b_v_SPC"
 /note="Organ: breast; Vector: pZeo-1; Site 1: SphI;
 Site 2: SphI; Ductal carcinoma in situ, high-grade, comedo
 , from 41 yo female. Library constructed in the laboratory
 of K. Polyak, M.D., Ph.D. (Dana-Farber Cancer Institute)."
 BASE COUNT 138 a 186 c 154 g 125 t 3 others

ORIGIN
 Query Match 22.8%; Score 585.4; DB 9; Length 606;
 Best Local Similarity 97.7%; Pred. No. 3.1e-81;
 Matches 592; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

880 GGTGAGCTGATCATCTACCGGCGCCACCTTCCCTCCAGGAGGAGGAGGAGGAGGAG 939
 Db 1 GGTGAGCTGATCATCTACCGGCGCCACCTTCCCTCCAGGAGGAGGAGGAGGAGGAG 60
 940 TACTCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 999
 Db 61 TACTCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
 1000 GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1059
 Db 121 GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
 1060 AACTGGAAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1119
 Db 181 AACTGGAAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
 1120 CTTCTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1179
 Db 241 CTTCTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300

Tue Apr 1 06:01:26 2003

us-09-988-971-1.rst

Page 12

```
Qy 1180 GTCTCTTGGATGATGCTAGGCCCCAAAGAGAGAGCCAAAAGGAAACCAAGGCTGCACA 1239
    |||||
Db 301 GTCTCTTGGATGATGCTAGGCCCCAAAGAGAGAGCCAAAAGGAAACCAAGGCTGCACA 360
    |||||
Qy 1240 CCTAGAACCCCAATTGAGCTCTTGGGCACTCCAGAGGCAAGCTGTGCACTCAGGAGG 1239
    |||||
Db 361 CCTAGAACCCCAATTGAGCTCTTGGGCACTCCAGAGGCAAGCTGTGCACTCAGGAGG 420
    |||||
Qy 1300 GAGGTTGGGACACAGAGGTGCACTAGGGTCCCACTGTACCTTGTCTTCTCTCTT 1359
    |||||
Db 421 GAGGTTGGGACACAGAGGTGCACTAGGGTCCCACTGTACCTTGTCTTCTCTCTT 480
    |||||
Qy 1360 AGCCCTTAGAAGTACCTACTCTCTCCAGTGCATGATCCCACTGGGACCTCTAGTGC 1419
    |||||
Db 481 AGCCCTTAGAAGTACCTACTCTCTCTCCAGTGCATGATCCCACTGGGACCTCTAGTGC 540
    |||||
Qy 1420 GAGTCAGAGAGGTTGGGACACAGGCCCCAGGGTTCCAAAAAGAAATTAAGCTCTGGGGG 1479
    |||||
Db 541 GAGTCAGAGAGGTTGGGACACAGGCCCCAGGGTTCCAAAAAGAAATTAAGCTCTGGGGG 600
    |||||
Qy 1480 GTCTGA 1485
    |||||
Db 601 GTCTGA 606
    |||||
```

Search completed: March 30, 2003, 02:49:10
Job time : 5065.55 secs